PAPER AAS 93-169

Higher-Order Ionospheric Effects on the GPS Observable and Means of Modeling Them

Sassan Bassiri
George A. Hajj

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AAS/AIAA Space flight Mechanics Meeting

PASADENA, CALIFORNIA FEBRUARY 22-24, 1993

AAS Publications Office, P. O. Box 28130, San Diego, CA 92198

Higher-Order Ionospheric Effects on the GPS Observable and Means of Modeling Them

Sassan Bassiri
Sharif University of Technology, Tehran, Iran

George A. Hajj
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Based on realistic modeling of the electron density of the ionosphere and using a dipole moment approximation for the earth magnetic field, we are able to estimate the effect of the ionosphere on the GPS signal for a ground user, The lowest-order $(1/f^2)$ effect, which is of the order of ,1-30 meters of zenith group delay, is subtracted out by forming a linear combination of the dual frequencies of the GPS signal. One is left with second- $(1/f^3)$ and third-order $(1/f^4)$ effects which are estimated typically to be -O-2 cm, and ~0-2 mm at zenith respectively, depending on the time of day, time of year, the solar cycle and the relative geometry of the magnetic field and the line of sight. Given the total electron content along a line of sight, we derive an approximation to the second-order term which is accurate to -90% within the magnetic dipole moment model; this approximation can be used to reduce the second-order term to the millimeter level, thus, potentially improving precise positioning in space and on the ground, The induced group delay, or phase advance, due to second- and third-order effects are examined for two ground receivers located at equatorial and mid-latitude regions tracking severa I GPS satellites.

1. Introduction

The Global Positioning System (GPS) consists of 24 satellites, evenly distributed in 6 orbital planes around the globe, at about 20200 Km altitude, Precise positioning of the GPS satellites as well as ground and space users is now reaching few parts in 10^9 (Yunck et al., 1985; Lichten and Border, 1987; Yunck et al., 1987; Wu et al. 1991; Lichten, Mar. 1990; Lichten, Apr., 1990), In addition, the GPS has been heavily utilized in a host of geodetic and other applications. These include seismic tectonic motions (Thornton et al., 1986; Kellogg and Dixon, 1990; Freymueller and Kellogg, 1990), earth orientation studies (Malla and Wu, 1989; Freedman, 1991), gravimetry (Bertiger et al., 1992), atmospheric water vapor calibration (Lichten, Feb. 1990) and ionospheric monitoring (Coco, 1991), Precise positioning and other GPS-based applications however require a very good

understanding of all effects on the GPS signal as it propagates through the earth atmosphere, so that they can be solved for or modeled.

The GPS transmits two right-hand circularly polarized (RCP) signals, at L band frequencies, L1 at 1574.42 MHz and 1..2 at 1227.6 MHz, which correspond to wavelengths of 19.0 cm and 24,4 cm respectively, These are modulated by a pseudo-random Precision code (P-code) at a frequency of 10.23 MHz (Spilker, 1978). (The additional lower frequency C/A modulation does not concern us here.) A single measurement, for a given transmitter/receiver pair, will consist of four observable which we will denote by L_1, L_2 for the accumulated carrier phase measurements at the two frequencies and P_1, P_2 for the corresponding P-code pseudorange. In addition to the geometric range delay, the signals will experience delays, or phase advances, due to the presence of the ionosphere and neutral atmosphere.

The delay due to the neutral atmosphere is nondispersive and therefore is the same for all observable, its effect is of the order of 2 meters and can be solved for to better than a centimeter (Lichten, Feb. 1990; Tralli and Lichten). However, due to the dispersive nature of the ionosphere, the group delay (or phase advance) caused by it is frequency dependent, and is of the order of.1 -30 meters at zenith depending on the time of day, time of year and the solar cycle, If the ionospheric effect on signal delay (or phase advance) is expanded in powers of inverse frequencies, then the lowest order term ($1/f^2$ term), by far the most dominant, can be solved for and subtracted out by virtue of the dual frequencies of GPS. Higher-order terms ($1/f^3$ and $1/f^4$ terms, referred to as second- and third-order terms in the following analysis) are of the order of sub-millimeters to several centimeters, which remain embedded in the signal and contribute to range and accumulated phase errors. While the first-order term depends simply on the Total Electron Content (TEC), namely the integrated electron density inside of a columnar cylinder of unit area between the transmitter and the receiver, higher-order terms depend on the coupling between the earth magnetic field and the electron density everywhere along the line of sight, In order to estimate the higher-order effects on the GPS observable, we model the ionosphere by a sum of Chapman layers and the earth magnetic field by that of a dipole moment. Such a model will allow us to estimate the effect at different geographical locations on the ground as well as its sensitivity to the electron density distribution, It will be demonstrated that knowledge of the TEC can be used to calibrate most of the second-order effect and reduce P-code and phase measurement errors to sub-centimeter.

Due to the inhomogeneity of the propagation medium, the GPS signal does not travel along a perfectly straight line. Moreover, since the medium is dispersive, the two frequencies will take two slightly different paths, Applying the empirical formula given by Brunner and Gu (1991) on the ionospheric mode] used below, the residual range error, which is the difference between the dual-frequency corrected range and the true range, due. to bending alone, is estimated to be -4 mm at 10^o elevation angle and less than a millimeter for elevations above 30^o. The bending effect will be ignored in the following analysis; the two signals will be assumed to travel along the same straight path.

A more elaborate modeling of higher-order ionospheric effects, where bending is taken into account, have been considered by Brunner and Gu (1991; see also Gu and Brunner, 1990). In their paper the International Geomagnetic Reference Fields and a Chapman

profile of the ionosphere were used to estimate the residual range error, They also proposed an improved linear combination that corrects for the second- and third-order terms as well as bending. Their improved linear combination requires knowledge of N_m and h_m, the electron density peak and its altitude respectively. In this paper, the second- and third-order terms are considered separately. We estimate that the second-order term is dominant most of the time over the third-order and the curvature terms. We suggest a method of modeling the second-order effect based on a thin shell model of the ionosphere and a dipole magnetic field. The second- and third-order errors are considered at different geographical locations, tracking different satellites. It is demon.. trated that knowledge of the TEC alone can be used to reduce the higher-order effects to few millimeters. All throughout this manuscript the MKS system of units is used unless otherwise stated.

2. Propagation of EM Waves in the Ionosphere

The ionosphere is a microscopically neutral ionized gas consisting principally of free electrons, ions, and neutral atoms or molecules. Positive ions are 2,000 to 60,000 times more massive than an electron. Thus at the frequencies used for radio communication, the range of movement of an ion caused by the electric field of a radio wave is smaller than that of an electron by about the same factor (Budden, 1985). This implies that the effect of ions on the propagation of radio waves at GPS frequencies can be ignored .

When a magnetostatic field B_o is applied to a plasma, the plasma becomes anisotropic for the propagation of electromagnetic waves. That is, the scalar dielectric constant of the plasma is transformed into a tensor, To study the propagation and polarization properties of a plane monochromatic wave in a magnetically biased homogeneous lossless plasma we regard the plasma as a continuous medium whose conductivity is zero, whose permeability is equal to that of vacuum, and whose dielectric constant is a tensor. By solving the Helmholtz wave equation subject to proper constitutive relations, one can obtain the expressions for fields and for the index of refraction. The index of refraction, n, for the earth's ionosphere is given by the Appleton -Hartree formula (Papas, 1965) as follows:

$$n_{\pm}^2 = 1 - \frac{2X(1-X)}{2(1-X) - Y_{\perp}^2 \pm \sqrt{Y_{\perp}^4 + 4(1-X)^2 Y_{\parallel}^2}}, \qquad (1)$$

where

$$x = \left(\frac{f_p}{f}\right)^2 = \frac{(Ne^2/4\pi^2\varepsilon_o m)}{f^2}, \qquad (2)$$

$$Y_{\perp} = Y \sin\theta_B \; ; \; Y_{\parallel} = Y \cos\theta_B, \qquad (3)$$

$$Y = \left(\frac{f_g}{f}\right) = \frac{(|e| B_o/2\pi m)}{f} \qquad (4)$$

N is the number density of electrons, e and m are the electron charge and mass, respectively, ε_0 is the permittivity of the free space, f_p, f_g and f are the plasma-, gyro- and carrier frequencies, respectively, θ_B is the angle between the earth's magnetic field B_o and the direction of propagation of the wavefront k. B_o is the magnitude of B_o. By definition $Y = eB_o/2\pi fm$, and since for electrons e is negative, Y is antiparallel to B_o. The plasma-frequency is the natural frequency of oscillation for a slab of neutral plasma with density N after the electrons have been displaced from the ions and are allowed to move freely. The gyro-frequency is the natural frequency at which free electrons circle around the magnetic field lines, For the earth's ionosphere, with N = 10^{12} electrons/m^3 the plasma-frequency $f_p \approx 8.9$ MHz. The gyro-frequency for an electron in earth's magnetic field (--2x1 O-5 Tesla) is $f_g \approx .59$ MHz.

The plus and minus signs of Eq. (1) correspond to the ordinary and extraordinary wave modes of propagation respectively. In general, these two waves are elliptically polarized with left and right senses of rotation, respectively. As a result of different phase velocities of the two waves, the total wave (sum of ordinary and extraordinary waves) undergoes Faraday rotation as it passes through the ionosphere. When the carrier frequency is large compared to plasma- and gyro-frequencies, the principle modes of propagation are dominantly circularly polarized, This is the case for the GPS carrier frequencies.

Under the assumption that $Y << 2|\cos\theta_B|(1-X)/\sin^2\theta_B$, the index of refraction can be expanded in inverse powers of frequency. For the GPS carrier frequencies we have (f_p/f) = 5.65 x 10^{-3} and 7.25x 10^{-3}, as well as (f_g/f) = 3.75x10^{-4} and 4.81x10^{-4} for L1 and L2 respectively. Therefore, the stated assumption is valid for GPS frequencies up to a value of $\theta_B \approx 890$. The expansion of Eq. (1) up to the fourth inverse powers of frequency gives

$$n_{\pm} = 1 - \frac{1}{2}X \pm \frac{1}{2}XY|\cos\theta_B| - \frac{1}{4}X\left[\frac{1}{2}X + Y^2(1 + \cos^2\theta_B)\right]. \qquad (5a)$$

The second, third, and fourth terms on the right-hand side of Eq. (5a) are proportional to the inverse square, inverse cube, and inverse quartic powers of frequency, respectively. The two values of n refer to the ordinary (+) and extraordinary (–) waves. In the special case where $\theta_B = 90^o$, $Y_{\|}$ vanishes and Eq. (1) reduces to

$$n_{\pm}\Big|_{(\theta_B=\frac{\pi}{2})} = 1 - \frac{1}{2}X - \frac{1}{8}X^2 - \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} X\,Y^2 + O\left(\frac{1}{f^6}\right). \qquad (5.b)$$

The upper and lower coefficients of XY^2 correspond to the upper and lower signs of $n_{\pm}$ respectively. If we compare Eq. (5a) in the limit $\theta_B \to 90^o$ with Eq. (5. b), the following observations can be made, Both equations give a vanishing contribution to the $1/f^3$ term; both have the same coefficient of X^2; they differ in the coefficient of XY^2. However, from order of magnitude considerations made above, the term XY^2 is two order of magnitude smaller than X^2 and can be neglected. Therefore, we will consider Eq. (5a) to be generally valid for all values of θ_B in the following discussion.

At this point it should be noticed from Eq. (5a) that the index of refraction is smaller than unity, which corresponds to a phase velocity greater than the speed of light in the vacuum (phase advance), The group refractive index, on the other hand, given by $n_{\pm}^{group} = n_{\pm} + f(dn_{\pm}/df)$, can be written as

$$n_{\pm}^{group} = 1 + \frac{1}{2}X \mp XY|\cos\theta_B| + \frac{3}{4}X\left[\frac{1}{2}X + Y^2(1+\cos^2\theta_B)\right] . \quad (6)$$

The group delay of a signal passing through the ionosphere, relative to vacuum as a reference, can be rewritten as

$$\tau_{\pm}^{group} = \frac{1}{c}\int (n_{\pm}^{group}\cos\alpha - 1)\, d\ell \quad , \quad (7)$$

where $d\ell$ is an element of length along the line of sight, c is the velocity of light in vacuum and *a is* the angle between the wave normal and the ray direction. This angle has significance in anisotropic media, where the direction of the wave normal is, in general, different from the direction of energy propagation. Angle *a* can be found from the following relation: tana = $(1/n)\,\delta n/\delta\theta_B$. Using Eq. (5a) and the definition of *a,* it is easy to show that for the GPS carrier frequencies *cosa is* essentially unity. Using Eqs. (5a), (6) and (7) the GPS observable can be written, in the dimension of length, as

$$P_1 = \rho + \frac{q}{f_1^2} + \frac{s}{f_1^3} + \frac{r}{f_1^4}, \quad (8a)$$

$$P_2 = \rho + \frac{q}{f_2^2} + \frac{s}{f_2^3} + \frac{r}{f_2^4}, \quad (8.b)$$

$$L_1 = \rho + n_1\lambda_1 - \frac{q}{f_1^2} - \frac{1}{2}\frac{s}{f_1^3} - \frac{1}{3}\frac{r}{f_1^4}, \quad (9a)$$

$$L_2 = \rho + n_2\lambda_2 - \frac{q}{f_2^2} - \frac{1}{2}\frac{s}{f_2^3} - \frac{1}{3}\frac{r}{f_2^4}, \quad (9.b)$$

where

$$q = \frac{1}{2}\int f_p^2 d\ell = (40.3)\int N\, d\ell = 40.3\ TEC, \quad (1o)$$

$$s = (\mp)\int f_g f_p^2 |\cos\theta_B| \, dl = (\mp 7527)\, c \int NB_o |\cos\theta_B| \, dl \,, \qquad (11)$$

$$r = (2437)\int N^2 \, dl + 4.74 \times 10^{22} \int NB_o^2 \left(1 + \cos^2\theta_B\right) dl \,. \qquad (12)$$

TEC is the total electron content along the line of sight, and *A is* the operating wavelength, In Eqs. (8) and (9), *p* corresponds to the geometrical distance plus all the nondispersive terms that are common to both frequencies such as clocks, transmitter and receiver delays and the neutral atmospheric delay, In Eq. (9) $n_1\lambda_1$ and $n_2\lambda_2$ correspond to unknown integer numbers of cycles that are constants for a given transmitter/receiver pair over a continuous tracking period. In addition to the terms shown on the right hand side of Eqs. (8) and (9), there are terms due to multipath, thermal noise, phase center variations and a transmitter/receiver relative geometry dependent term; however, these are not the. subject of our study here, and are omitted from Eqs. (8) and (9). The (-) and (+) signs in Eq. (11) correspond to the ordinary and extraordinary waves respectively. Ignoring the LCP component of the GPS signal which has less than ,35% and 2.5% of the total power, for L1 and L2 respectively, only the (+) sign will be of relevance to us in the subsequent analysis.

3. Ionospheric Layers and Geomagnetic Field Models

To proceed with the computation of the higher-order delays, we have to assume models for the electron density, N, and the earth magnetic field B_o.

The bulk of the earth ionosphere extends from about 80 to 1000 km altitudes. The ionospheric electron density profile exhibits several distinct overlapping layers (D, E, F1, F2), as a result of the competing processes of particle production, loss and transport, Each of these layers has a maximum density at a certain altitude, above and below which the ionization density tends to drop off. The maximum electron densities (10^{12} to 10^{13} m^{-3}) are observed at the F2 peak; the peak altitude ranges from 250 to 350 km at mid-latitudes and from 350 to 500 km at equatorial latitudes. 'I'he F1 region, which is present during the day but absent at night, has a peak near 200 km altitude and is 3-5 times smaller than that of F2. The E peak density is about one order of magnitude smaller than the F2 peak and typically located at 100 to 120 km altitude. During daytime there is also a D layer with a peak at 80 km altitude (Jursa, 1985).

In order to estimate the higher order ionospheric effects on electromagnetic signal propagation we model the ionosphere as the sum of three Chapman layers as shown in figure 1. The Chapman layer was discussed in detail by Rishbeth and Garriott (1969), It is derived by assuming a homogeneous composition for air at a constant temperature, The

curvature of earth is neglected and it is assumed that the atmosphere is horizontally stratified and the scale height H_s *is* independent of height. As the solar radiation travels downward through the atmosphere, it is absorbed and hence ionization is produced. The rate of electron production is a function of height above mean sea level h and the sun's zenith angle χ which is the angle between the ray from the sun and the zenith, From considerations of production of electrons by photoionization and their removal by recombination, the following formula for the electron density distribution can be obtained (Rishbeth and Garriott, 1969):

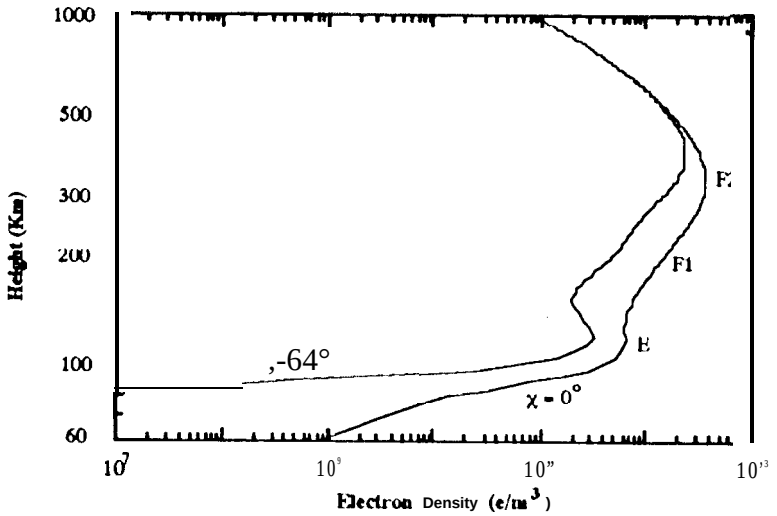

Figure 1: Ionospheric profile modeled as the sum of three Chapman layers.

$$N = N_{max} \exp\left[\frac{1}{2}(1 - z - e^{-z}\sec\chi)\right], \qquad (13)$$

where N_{max} is the maximum value of the electron density at an altitude of h_{max} and $z = (h-h_{max}) / H_s$. When χ is near *900,* as near sunrise and sunset, the plane earth approximation fails. To correct for this, sec χ in Eq. (13) is replaced with the grazing incidence function $Ch(x,\chi)$. This function, which applies accurately only to a spherically symmetric atmosphere with H_s independent of height, can be expressed as

$$Ch(x,\chi) = \left(\frac{1}{2}\pi x \sin\chi\right)^{1/2} e^{\frac{1}{2}x\cos^2\chi} \times \left[1 \pm \operatorname{erf}\left(\frac{1}{2}x\cos^2\chi\right)^{1/2}\right], \qquad (14)$$

where $x = (R_E + h)/H_s$, R_E is the earth's radius, and erf(.) is the error function. The plus (minus) sign refers to $\chi > 90^o$ ($\chi < 900$). Figure 1 is a plot of the electron density distribution versus height for two different solar zenith angles $\chi = 0^o, 64^o$. In obtaining this distribution the three different Chapman layers that are added together resemble the

ionospheric F_2, F_1 and E layers, and have the following parameters (obtained from (Jursa, 1985))

E layer: h_{max} =110 km, H_s =11 km, N_{max} = .5x10^{12} (m^{-2});
F1 Layer: h_{max} =210 km, H_s =52 km, N_{max} = 1.0x10^{12} (m^{-2});
F2 Layer: h_{max} =350 km, H_s =78 km, N_{max} = 3.25x1012 (m^{-2}).

These numbers are representative of a day time profile typical of a year near sunspot maximum. The D layer which is normally present during the day time is not included. During night time, the F1 layer disappears and the electron density for a given height are about 10-100 times smaller than that of day time. In a solar minimum the same features (D,E,F1,F2 layers) are preserved with the electron density scaled down roughly by a factor of 10.

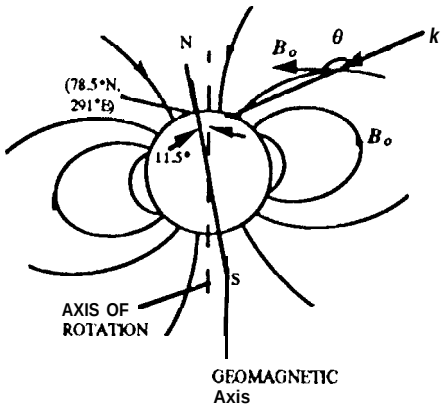

Figure 2: The earth's magnetic field modeled as an earth-centered dipole, aligned along the geomagnetic axis,

Next we need to model the earth's magnetic field. A first approximation to the geomagnetic field near the surface of the earth is an earth-centered dipole with its axis tilted to intersect the earth at 78.5°N latitude, 291 .0°E longitude, which corresponds to the geomagnetic north pole; and at 78.5°S latitude, 111.0°E longitude which corresponds to the geomagnetic south pole (Jursa, 1985) (see Fig. 2).

At this point we distinguish between two reference frames with a common origin at the earth's center, The geodetic frame is earth fixed and is given by $\hat{x},\hat{y},\hat{z}$ where $\hat{z}$ is along the earth's spin axis, and $\hat{x}$ is pointing toward O longitude. The geomagnetic frame, on the other hand, is obtained by first rotating the geodetic frame by an angle β = 2910 around its $\hat{z}$ axis, and then applying a second rotation by an angle δ =11.50 around the new $\hat{y}_m$ axis (Fig, 3). This geomagnetic frame is denoted by $\hat{x}_m,\hat{y}_m,\hat{z}_m$ and is constructed so that $\hat{z}_m$ is

along the magnetic dipole. A vector transformation from the geodetic to the geomagnetic frame is given by

$$V_m = \begin{pmatrix} \cos\delta \cos\beta & \cos \& \sin\beta & -\sin\delta \\ -\sin\beta & \cos\beta & 0 \\ \sin\delta\cos\beta & \sin\delta\sin\beta & \cos\delta \end{pmatrix} V \quad . \qquad (15)$$

At a point on the earth's surface we denote local geodetic east, north and vertical by $\hat{X},\hat{Y},\hat{Z}$, and geomagnetic east, north and vertical by $\hat{X}_m,\hat{Y}_m,\hat{Z}_m$ (Fig. 3). The magnetic field vector is given by

$$B_o = B_g\left(\frac{R_E}{r_m}\right)^3 \sin\theta_m \hat{Y}_m - 2B_g\left(\frac{R_E}{r_m}\right)^3 \cos\theta_m \hat{Z}_m, \qquad (16)$$

where r_m is the radial distance, and θ_m is the magnetic colatitude. B_g is the amplitude of the magnetic field at the earth's surface, at the magnetic equator, and is approximately equal to 3.12x1 O-5 Tesla.

4. Ananlysis

First-order effect

According to Eqs. (8) and (10), the first-order ionospheric delay can be written as 4.48x $10^{-16}\lambda^2 TEC$ (meters). For the GPS L1 and L2 frequencies, respectively, this translates to 16.2 cm and 26.7 cm of group delay (or phase advance) for every one TEC unit. (1 TEC unit = 10^{16} electron/m^2.) Zenith TEC measurements from the ground to GPS vary between 1-100 TEC units, depending on the time of day, the season and the solar activity amongst other factors. Therefore, first-order zenith ionospheric group delay (phase advance) ranges between ~.1-16 meters for L1 and -.2-27 meters for L2.

The first-order ionospheric term, which is about three orders of magnitude larger than higher-order terms, can be eliminated by using the "ionospheric free" linear combination, which, based on Eq. (8), is given by

$$\left(\frac{f_1^2}{f_1^2 - f_2^2}\right)P_1 - \left(\frac{f_2^2}{f_1^2 - f_2^2}\right)P_2 = \rho - \frac{s}{f_1 f_2 (f_2 + f_1)} - \frac{r}{f_1^2 f_2^2} \quad . \qquad (17.a)$$

A similar linear combination to Eq. (17.a) can be obtained using the accumulated phase. measurements of Eq. (9) to obtain

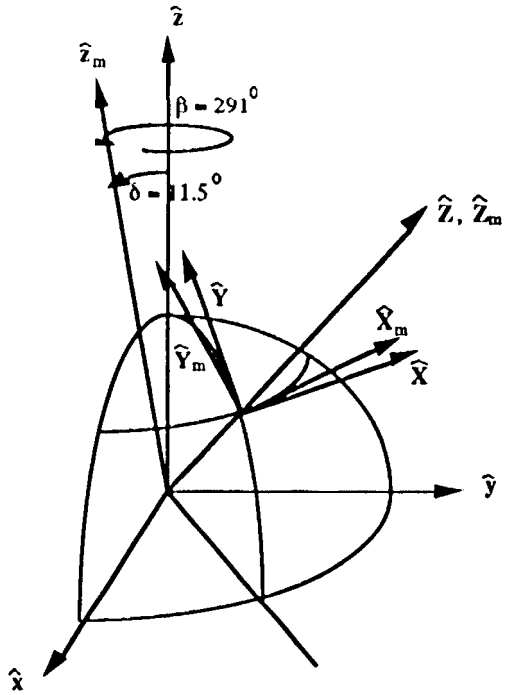

Figure 3: A graphical illustration of all the frames used in the text. $\hat{x},\hat{y},\hat{z}$ *correspond to the geodetic frame,* $\hat{x}_m,\hat{y}_m,\hat{z}_m$ *to the geomagnetic frame,* $\hat{X},\hat{Y},\hat{Z}$ *to geodetic local east, north and vertical, and* $\hat{X}_m,\hat{Y}_m,\hat{Z}_m$ *to geomagnetic local east, north and vertical*

$$\left(\frac{f_1^2}{f_1^2 - f_2^2}\right)L_1 - \left(\frac{f_2^2}{f_1^2 - f_2^2}\right)L_2 = \rho + \frac{s}{2f_1f_2(f_2 + f_1)} + \frac{r}{3f_1^2f_2^2} + B \quad , \quad (17.b)$$

where *B* corresponds to a phase bias that is constant for a given transmitter/receiver pair over a continuous tracking period. In the literature, the additional terms to the range *p on* the RHS of Eq. (17.b) (not including the fixed bias term) are referred to as Residual Range Error (RRE). In our study below we will refer to the (Residual Range Error)$_{group}$, $(RRE)_g$, namely the difference between the dual-frequency corrected range and the true range as inferred from the group delay measurements. The $(RRE)_g$ corresponds to the 2nd and 3rd terms on the RHS of Eq. (17.a).

As the first-order ionospheric term is eliminated, the dominant ionospheric errors are due to the second- and third-order terms which are discussed below.

Second-Order Effect

The term $B_o|\cos\theta_B|$ in Eq. (11) represents the absolute value of the component of the B_o field along the line of propagation; therefore, it can be replaced by $|\vec{B}_o \cdot \hat{k}|$, where $(\cdot)$ represents the inner product, and $\hat{k}$ is the unit vector in the direction of propagation.

Consider a station with magnetic colatitude and longitude θ_m and ϕ_m respectively, observing a satellite with elevation E_m and azimuth A_m, where A_m is measured from magnetic north. Then $\hat{k}$ is given by

$$\hat{k} = -(\cos E_m \sin A_m \hat{X}_m + \cos E_m \cos A_m \hat{Y}_m + \sin E_m \hat{Z}_m); \quad (18)$$

therefore,

$$|B_o \cdot \hat{k}| = B_g \left(\frac{R_E}{r_m}\right)^3 \times \left| \sin\theta_m' \cos E_m \cos A_m - 2 \cos\theta_m' \sin E_m \right|, \quad (19)$$

where θ_m, r_m are the magnetic colatitude and radial distance of a point along the link, respectively. This term, multiplied by the electron density, is the integrand of Eq. (1 1), where we must think of r_m and θ_m as varying along the line of integration, While the exact distribution of electron density along the line of sight is needed to calculate the second-order delay term, a useful approximation can be derived by assuming the ionosphere to consist of a very thin layer at altitude H. Then, the corresponding r_m and θ_m at the intersection point between the line of sight and the ionospheric layer are given by (for $Em > 10^o$)

$$r_m = R_E + H; \quad (20\text{-a})$$

$$\theta_m' = \theta_m - \frac{H}{R_E \sin E_m} \cos A_m \cos E_m + O\left(\frac{H^2}{R_E^2}\right). \quad (20\text{-b})$$

Combining Eqs. (8),() 1),(19) we can approximate the second-order ionospheric group delay (in units of distance) by

$$2^{rd} \text{ order ion, group delay} = 2.61 \times 10^{-18} \lambda \left(\frac{R_E}{r_m}\right)^3 \times \left| \sin\theta_m' \cos E_m \cos A_m - 2 \cos\theta_m' \sin E_m \right| TEC, \quad (21)$$

where r_m and θ_m' are given by Eq. (20). Setting H at 300 km, and ignoring the factor between the absolute signs, Eq. (21) implies that, in the dipole approximation, the second-order ionospheric group delay is of the order of .16 mm and .33 mm for L1 and L2 respectively, for each TEC unit. The second-order ionospheric phase advance, on the other hand, is one-half of this effect, When taking the ionospheric free linear combination, some cancellation in the second-order term takes place; the $(RRE)_g$ is then of the order of -.11 mm per TEC unit. The relations between the magnetic colatitude and longitude, θ_m and ϕ_m, and the geographical colatitude and longitude, θ and@, are given by

$$\text{Cos}\,\theta_m = \sin\delta\cos\beta\sin\theta\cos\phi + \\ + \sin\delta\sin\beta\sin\theta\sin\phi + \cos\delta\cos\theta\,, \tag{22}$$

$$\tan\phi_m = \frac{-\sin\beta\sin\theta\cos\phi + \cos\beta\sin\theta\sin\phi}{\cos\delta\,(\cos\beta\sin\theta\cos\phi + \sin\beta\sin\theta\sin\phi) - \sin\delta\cos\theta} \tag{23}$$

The satellite elevation in a local magnetic East-North-Vertical coordinates, E_m, is the same as the elevation in a local geodetic East-North-Vertical coordinates, E. On the other hand, the azimuths in these two coordinates are related through

$$A_m = A + \arccos(\sin\phi\sin\phi_m\cos\delta\cos\beta \\ + \cos\phi\cos\phi_m\cos\beta + \sin\phi\cos\phi_m\sin\beta \\ - \cos\phi\sin\phi_m\cos\delta\sin\beta)\,. \tag{24}$$

Figure 4 shows the absolute value of the $(RRE)_g$ due to the second-order term. This is shown for two stations at different longitudes and latitudes, tracking different GPS satellites, as indicated on the figure, These errors are calculated using the exact integral form of Eq. (1 1), and assuming the Chapman layer distribution of Fig. 1 and the magnetic field of a tilted dipole as described above. The angle χ in Eqs. (13) and (14) is determined based on the assumption that the $\hat{x}$ axis (Fig. 3) is pointing toward the sun at 12:00 universal time. The exact calculation, referred to as truth, is compared with an approximation obtained from Eqs. (20)-(24). According to the examples of Fig, 4, the true second-order absolute $(RRE)_g$ has an RMS value of 1.25 cm, and can be as large as 4 cm at the lowest elevation angle (10°). Using the thin layer model at 300 Km altitude as described above, it is possible to approximate this effect to better than 90% on the average. The difference between the truth and the approximation has an average of.11 cm and a variance of .25 cm. This suggests that a thin layer model of the ionosphere can be very useful in calibrating the second order ionospheric effect and therefore improving GPS range measurements.

It is worth mentioning at this point that a three frequency ranging system will add a third part to Eq. (8) and, therefore, allowing the removal of the second order term. This possibility is explored in some detail by Bassiri (1990).

Third-Order Effect

Upon examining Eq. (12), we find that the second-term, except during times of very strong magnetic storms, contributes no more than sub-millimeter of range error for GHz frequencies. Therefore, we turn our attention to the first term, which can be simplified to (in MKS units)

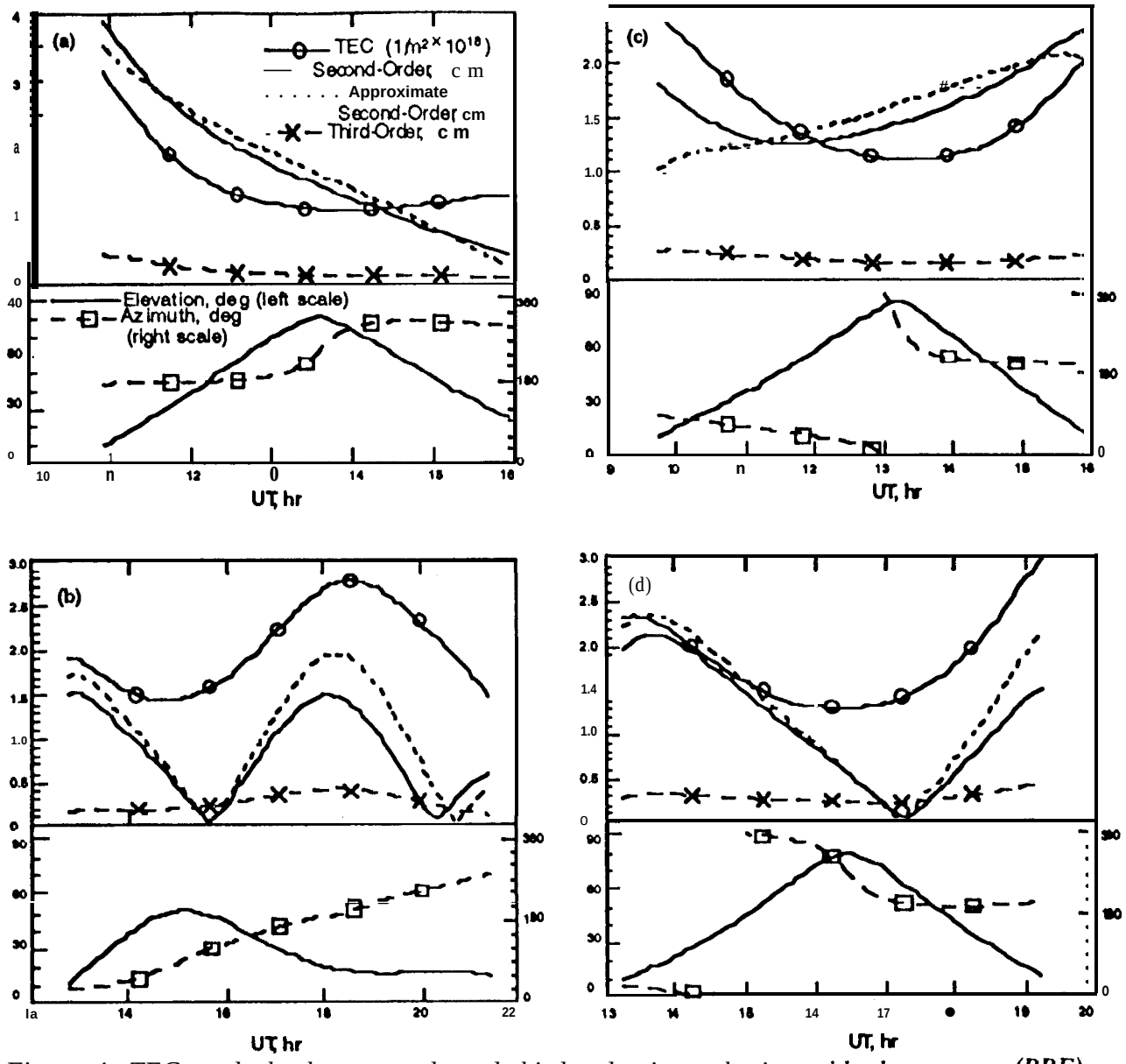

Figure 4: TEC, and absolute second- and third-order ionospheric residual range error, $(RRE)_g$, along the line of sight for different GPS-ground receiver. Shown also are the elevation and azimuth of the observed satellite as functions of time: (a) GPS 1, station at 40 deg N latitude, O&g longitude; (b) GPS 9, station at O deg latitude, 75 deg W longitude; (c) GPS 20, station at 40 deg N latitude, O deg longitude; and (d) GPS 16, station at O deg latitude, 74&g W longitude. The TEC values are calculated based on the Chapman layer of figure 1. The second-order $(RRE)_g$ is calculated based on two different methods. First: using Eq. (11) and the Chapman layer of Fig. 1 (solid line). Second: using Eqs. (20)-(24) (dashed line). The third-order term is calculated based on the Chapman layer model of Fig. 1 and Eq. (12).